International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4425 - 4429
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0578

Deepfake & Synthetic Media Detection

Tejaswini Lokhande[1], Shreya Ghadge[2], Janhavi Lakeri[3], Samruddhi Shedage[4], Snehal Pawar[5], Rutuja Pawar[6]
[1]Associate professor Dept of Computer Science & Engineering Yashoda Technical Campus, Satara-415015, Maharashtra, India.
[2,3,4,5,6]Student Dept of Computer Science & Engineering Yashoda Technical Campus, Satara-415015, Maharashtra, India.
Emails: *tejaswinilokhandeengg@yes.edu.in[1], shreyaghadge41@gmail.com[2], janhavilakeri@gmail.com[3], sshedage2004@gmail.com[4], pawarsnehal020@gmail.com[5], rutuupawar14@gmail.com[6]*

Abstract
Deepfake and synthetic media technologies have quickly changed with the growth of artificial intelligence. This has raised serious concerns about misinformation, security, and digital trust. This review paper looks at recent research on deepfake detection in image, video, and audio areas. It studies various methods like Convolutional Neural Networks (CNN), Generative Adversarial Networks (GAN), MesoNet, and Multilayer Perceptron (MLP) models to see how well they identify manipulated content. The review points out that while existing methods show high accuracy on controlled datasets, they struggle in real-world situations, including issues like compression, noise, and unfamiliar manipulation techniques. Most methods also focus on just one type of media, which limits their strength. The paper highlights crucial gaps, such as the lack of generalization, dependence on specific datasets, and limited integration of different types of media. Finally, it stresses the need for strong and flexible multimodal deepfake detection systems to tackle the growing challenges from synthetic media effectively.

Keywords *- Deepfake Detection ; Artificial Intelligence ; Convolutional Neural Networks (CNN) ; Mel-Frequency Cepstral Coefficients (MFCC) ; Multimodal Detection*

1. Introduction
Deepfake and synthetic media technologies have grown pretty quickly, mostly because of the growth of artificial intelligence AI, deep learning, and generative models. In general, these technologies can produce very convincing fake images video and even audio, and the result is that it becomes hard to tell what is real or not. Still, synthetic media is not only "bad", it also helps in entertainment, education, virtual reality, and digital content creation. But at the same time, it brings big problems like misinformation, cybercrime, identity theft, political manipulation, and broader digital security issues. And since tools for making deepfakes are now easier to access, people end up trusting digital information less, so there is this urgent need for detection systems that are actually reliable. To handle these problems, researchers have been using lots of deepfake detection approaches with machine learning and deep learning. You can see things like Convolutional Neural Networks (CNN), Generative Adversarial Networks (GAN), MesoNet, Multilayer Perceptron (MLP), and 3D Convolutional Neural Networks. In controlled settings, these methods have shown promising detection capability for manipulated media. However the deepfake generators themselves keep changing fast, so existing detection systems are still getting hit with major issues, even if they looked strong before. Even with some recent progress, today's deepfake detection techniques still come with limitations. Many models depend heavily on particular datasets, and they often fail to generalize when the manipulation type is new or unseen. Also, real-world conditions tend to lower performance, for example compression artifacts, low-quality media, changes in lighting, noise interference, and partial face occlusions. Plus, quite a lot of studies concentrate on just one modality, like image analysis or video analysis, while comparatively less attention is given overall

2. Literature Review
A bunch of researchers have been looking into how to detect deepfakes and synthetic media. They use all

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4425 - 4429
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0578

sorts of techniques, and it's clear there are some real steps forward but also a lot of hurdles still in the way. Powell and his team took a closer look at the technology behind synthetic content back in 2025. They explored the tools used to detect it and the policies in place globally. What they found was that deepfake technology is evolving at a rapid pace, far outstripping the efforts of governments and organizations to keep up. It's become clear that we need more comprehensive regulations that everyone can follow, as well as more advanced detection methods and education for people to understand the issue. However, their focus was primarily on the policy aspects, rather than the technical side of creating synthetic content. They highlighted the need for stronger guidelines and more effective ways to identify and mitigate the risks associated with deepfakes. By understanding the policy landscape, we can better address the challenges posed by this emerging technology. There's a research paper from IEEE Access in 2024 that put deepfake detectors to the test, specifically for use in things like court cases or investigations. The models performed pretty well when given clean, high-quality data sets. However, when they were faced with real-life complexities, such as videos that had been compressed or edited, their accuracy took a significant hit. This really highlights the fact that these systems aren't yet ready to be used as evidence in the real world. It's one thing to perform well in a controlled environment, but it's quite another to handle the messiness of real-life data. This is a big concern, because in real-world applications, like court cases, the stakes are high and the data is often far from perfect[1]. Researchers like Zia and their team made some progress in 2024, trying to improve synthetic media using GANs, for both creating and detecting it. They had some success with the data they were familiar with, but when it came to new or unknown types of GANs, their models just didn't cut it. It's like we need a detection system that can handle anything that comes its way, something more robust and versatile. We want to be able to catch any kind of synthetic media, no matter how it was made or what kind of GAN was used. The current methods are just not good enough, we need something better, something that can keep up with the constantly evolving landscape of synthetic media. Victoire and her team[2] discussed the issue of deepfakes being used maliciously, such as spreading false information or causing harm, in the same year. They emphasized that simply improving the technology itself is not a sufficient solution. Instead, a combination of legal measures, technical fixes, and raising public awareness is necessary to effectively address the problem. By taking a multi-faceted approach, it's possible to mitigate the negative consequences of deepfakes, but relying solely on technological advancements won't be enough to stop the issues. In 2024, Zia and colleagues took a closer look at detection methods based on GANs. What they found was that these methods work pretty well for the things they were trained on, but when it comes to new and clever ways of manipulating data, they can be thrown off. This suggests that we need to develop models that aren't tied to specific types of GANs, but can instead adapt to different situations. Researchers at Sun, back in 2023, made a breakthrough with a deep learning system that can detect deepfakes in videos with impressive accuracy and efficiency, without consuming too much computer power. However, this system still has its limitations - it struggles to identify deepfakes in low-quality videos or those that have been manipulated using techniques it wasn't trained on, which means generalization remains a challenge. This is a problem because the system needs to be able to adapt to new and unknown types of deepfakes in order to be truly effective. Despite this, the progress made so far is a significant step forward in the fight against deepfakes. In 2023, Goud and their team developed a model using MesoNet, focusing on the mid-level details of images. It performed reasonably well on standard test sets, but when it came to real-world data with more variation, it didn't quite live up to expectations. This was likely due to the fact that the training data lacked diversity, which made it difficult for the model to generalize effectively to new, unseen data. As a result, the model's performance suffered when faced with the complexities and nuances of real-world images. Going back to Lu et al. in 2021, they made this 3D attentional Inception CNN thing for videos. It grabs spatial and time based features pretty effectively. High scores on controlled setups, sure. But add in blocks or distortions like in real life, and

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4425 - 4429
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0578

performance drops. Robustness is the big question. Researchers like Das and others found a way to improve detection in 2021 by using dynamic face changes. This made models more robust and able to handle different types of data. However, deepfakes are evolving so quickly that it's challenging to keep up. We might need to come up with even more flexible methods to stay ahead. The field is moving fast, and right now, not all the pieces fit together perfectly[3].

3. Critical Analysis

When you look at the ways we currently detect deepfakes, you can see that they work really well in controlled situations, but they struggle when things get more real-world, like with compression or noise. A lot of these methods, especially the ones that use GANs, are great at spotting manipulations they've seen before, but they're not so good at handling new or unknown deepfake techniques. This is a big problem because it means these detection methods aren't flexible enough to keep up with the latest deepfake methods. Some research only looks at one type of media, like images or videos, which doesn't help when dealing with multiple types of data at the same time. There are models like MesoNet and 3D CNN that work really well in certain situations, but they can be thrown off by small changes or distortions in real-world data. Also, current methods often aren't strong enough and can't keep up with how fast deepfake technology is changing. This means they might not be able to detect new types of deepfakes, which is a big problem. We need approaches that can handle different types of media and adapt to new technologies quickly. One major issue with current systems is that they rely on huge amounts of high-quality data, which isn't always available when you need it. This means we need to create detection systems that are more flexible, robust, and can handle different types of data - and can do all this in the real world, not just in a lab. To make progress, researchers should work on making models that can adapt to new situations, use a wider range of datasets, and detect things in real-time. This would help these systems work better in everyday environments.

4. Key Features

- Use of Deep Learning Models Most papers use advanced models like CNN, GAN, MesoNet, and MLP for deepfake detection.
- GAN-Based Detection Techniques Several studies focus on Generative Adversarial Networks (GANs) for both generation and detection of synthetic media[4].
- High Accuracy on Controlled Datasets Many models achieve high performance on clean and well-structured datasets.
- Video Analysis using Spatial-Temporal Features Models like 3D CNN effectively capture spatial and temporal patterns in videos[5].
- Efficient Detection Models Some approaches focus on reducing computational cost while maintaining good accuracy.
- Face-Based Detection Methods Many techniques rely on face extraction and analysis for identifying deepfake content.
- Robustness Improvement Techniques Methods like data augmentation and feature-based learning are used to improve model performance[6].
- Consideration of Real-World Applications Some studies evaluate models for forensic and judicial use cases.
- Awareness of Social and Security Risks Research also highlights misinformation, cybercrime, and misuse of deepfake technology.

5. Limitations

- Poor Generalization to Unseen Data Models fail when tested on new or unseen deepfake techniques.
- Performance Drop in Real-World Conditions Accuracy decreases due to: Compression , Noise , Low-quality videos
- Dataset Dependency Models rely heavily on specific datasets, limiting real-world applicability.
- Single-Modality Focus Most research only looks at pictures or videos, and they forget about sound, which means they're not using all the ways to understand something.
- Sensitivity to Distortions Models struggle with: Occlusion , Lighting changes , Facial

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4425 - 4429
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0578

variations

- Rapid Evolution of Deepfake Techniques Existing models cannot keep up with new generation methods.
- Limited Real-World Testing Most models are tested only in controlled environments, not real-world scenarios.
- High Computational Requirements Some deep learning models require high processing power and resources.
- Lack of Standardization and Policies There isn't a standard set of rules or guidelines that everyone follows to detect deepfakes[7].

6. Results And Discussion

The performance of the proposed Deepfake and Synthetic Media Detection System was evaluated in terms of functionality, efficiency, accuracy, and usability across image, video, and audio detection modules.

6.1.Functional Evaluation

The proposed system successfully implements multimodal deepfake detection through separate modules for image, video, and audio analysis:

- The image module identifies manipulated facial features.
- The video module detects inconsistencies across extracted frames.
- The audio module classifies real and synthetic speech signals.

The integration of these modules into a unified framework demonstrates the effectiveness of automated multimodal deepfake detection.

6.2.Performance and Computational Efficiency

The system demonstrates efficient processing performance across all modules:

- The MTCNN algorithm provides fast and accurate face detection.
- EfficientNet-based models achieve high detection accuracy with good computational efficiency.
- Frame-based analysis reduces video processing complexity while maintaining reliable predictions.

Overall, the system provides efficient processing suitable for scalable and near real-time applications.

6.3.Accuracy and Evaluation Quality

The system produces reliable detection results across different media types:

- Image detection identifies visual artifacts in manipulated images.
- Video detection analyzes multiple frames for consistent prediction results.
- Audio detection uses MFCC features to distinguish between real and fake speech.

The multimodal approach improves overall detection reliability compared to single-modality systems.

6.4.Usability and User Experience

The system includes a user-friendly interface developed using HTML and CSS, allowing users to:

- Upload image, video, or audio files easily.
- View prediction results with confidence scores.
- Understand system outputs through a simple and transparent interface.

This improves accessibility for both technical and non-technical users[8].

6.5.Discussion

The results show that multimodal analysis improves the robustness and reliability of deepfake detection while reducing false predictions. However, some challenges still exist[9], including reduced accuracy for low-quality or compressed media, difficulty in detecting newly emerging deepfake techniques, and sensitivity of audio detection to background noise.

Conclusion

Based on what I've learned from researching how to detect deepfakes and synthetic media, it seems like deep learning techniques such as CNNs, GANs, and MLPs have made some real progress in identifying fake content. These methods are pretty good at catching visual or audio clues that don't seem quite right. For instance, they can pick up on tiny inconsistencies in a video or audio clip that might not be noticeable to the human eye or ear. This is a big deal, because it means we might be able to use these techniques to spot fake content before it spreads and causes harm So, I've been looking at these methods and they seem to work really well on the datasets they're tested on, you know, the ones that are all controlled and perfect. But, there are some problems that keep popping up. For example, they're not always

International Research Journal on Advanced Engineering Hub (IRJAEH)
e ISSN: 2584-2137
Vol. 04 Issue: 06 June 2026
Page No: 4425 - 4429
https://irjaeh.com
https://doi.org/10.47392/IRJAEH.2026.0578

able to handle new types of deepfakes that they haven't seen before, and when you use them in real-life situations, they're not as accurate as you'd like. I think the issue is that they're too reliant on those specific datasets, and that's just not how things work in the real world. So, most methods focus on just one kind of media, like video or audio, which can be a problem when you're dealing with something that combines both. And to make things worse, deepfakes are evolving really quickly, with new techniques being developed all the time, which keeps messing with the tools we use to detect them. This means that detection tools are constantly being thrown off, making it even harder to keep up with the latest deepfakes. It's like a cat-and-mouse game, where the deepfakes are always one step ahead, and we're struggling to catch up. So it seems like we need to come up with better systems that can handle a lot of different things and work well in the real world, which can be pretty messy. I'm probably making it sound simpler than it is, but maybe what we need to do is create models that can take a punch, use a wider range of data to train them, and find ways to catch problems as they're happening. The thing is, we're not really sure how to do that part yet, it's still a work in progress. We need systems that can adapt to different media and situations, and that can keep up with the pace of the real world. It's a tough problem to solve, but maybe if we can figure out how to make our models more robust and flexible, we can make some real progress.

Acknowledgements

I would like to express my sincere gratitude to my project guide Tejaswini Lokhande and faculty members of the Department of Computer Science and Engineering at Yashoda Technical Campus Satara for their valuable guidance and support during the completion of this review paper on deepfake and synthetic media detection. I also thank my friends and classmates for their suggestions and encouragement throughout the research work.

References

[1]. Loreen Powell, Hayden Wimmer, Carl Rebman (2025) Exploration of AI synthetic media and deepfake: understanding the technologies, detection software, legislation, initiatives, and curriculum
[2]. IEEE Access authors (2024) Empirical Assessment of Deepfake Detection: Advancing Judicial Evidence Verification Through AI
[3]. Rabbia Zia, Mariam Rehman, Afzaal Hussain, Shahbaz Nazeer and Maria Anjum (2024) Improving synthetic media generation and detection using generative adversarial networks
[4]. Dr. T. Amalraj Victoire, M. Vasuki, A. Aravindhan (2024) A Study on Deepfake and Synthetic Media: Combining Misinformation and Malicious Use
[5]. Rabbia Zia, Mariam Rehman, Afzaal Hussain, Shahbaz Nazeer, Maria Anjum (2024) Improving synthetic media generation and detection using generative adversarial networks
[6]. Ruipeng Sun, Ziyuan Zhao, Li Shen, Zeng Zeng, Yuxin Li, Bharadwaj Veeravalli, Xulei Yang (2023) An Efficient Deep Video Model for Deepfake Detection
[7]. Mrs. N. Swapna Goud A. Manoj Kumar, G. Bhanu Prakash, K. Mithil Reddy (2023) Synthetic Media Detection
[8]. Changlei Lu, Bin Liu, Wenbo Zhou, Qi Chu, Nenghai Yu (2021) Deepfake Video Detection Using 3D-Attentional Inception Convolutional Neural Network
[9]. Sowmen Das, Selim Seferbekov, Arup Datta, Md Saiful Islam, Md Ruhul Amin (2021) Towards Solving the DeepFake Problem: Improving DeepFake Detection Using Dynamic Face Augmentation